Incorporating Correlated Station Dependent Noise Improves VLBI Estimates

John M. Gipson
NVI, Inc./NASA GSFC, Code 698, Greenbelt, 20771 MD, USA

Abstract. I studied the effects of including station dependent delay noise in the analysis of geodetic VLBI data. Such terms increase the observational noise, and introduce correlation between observations. Using the CONT05 sessions, I demonstrate that introducing such noise terms reduces baseline scatter, gives more realistic formal errors, and improves agreement between VLBI and GPS estimates of Polar Motion and LOD.

1 Introduction

The VLBI observable, colloquially called the "delay", is the difference in arrival time of a signal at two different stations. Roughly speaking, the delay is measured by cross-correlating the signals received at the two stations and searching for a peak. This process has an uncertainty associated with it which, for clarity, I call *measurement* noise and denote by σ_{meas}. This noise is inherent to the correlation process and depends only on the signal strength, sensitivity of the antennas, frequency sequence, and number of bits recorded. One can show, assuming SNRs commonly used in VLBI, that the measurement noise on different baselines is uncorrelated. The standard assumption in VLBI processing is that the *observational* noise in VLBI is just the measurement noise. This has the corrolary that the VLBI observations are independent.

There are several indications that this assumption is false. First, χ^2 from individual session solutions is much larger than it should be if $\sigma_{meas} = \sigma_{obs}$. Second, baseline length scatter is larger than it should be based on the formal errors. Third, comparison of EOP measurements from simultaneous VLBI sessions are inconsistent with the formal errors. All of these are indicative of unmodeled error sources and/or incorrect modeling.

There are many other error sources besides measurement noise: 1) Phase cal errors; 2) RF interference in the signals; 3) Other correlator related errors; 4) Source structure; 5) Source position errors; 6) Errors in geophysical models; 7) Mismodeling clocks and/or atmospheres; 8) Underparametrizing the time variation of clocks and/or atmospheres; etc. All of these increase the noise of individual observations. Many also introduce correlation between observations.

In this note, I look at the special case of error sources which are manifest as station dependent changes in the measured delay. Examples are cable calibration errors, and atmosphere calibration or modeling errors. Since, at any instant, these errors are the same for all observations involving a station, these observations are not independent, and the observation covariance matrix is no longer diagonal. Neglecting these correlations leads to formal errors which are too small, and incorrect estimates of parameters.

Other scientists have studied the stochastic model used in VLBI. In a prescient paper Qian (1985) discussed the effect of correlation between observations, suggested a method for estimating these correlations empirically, and demonstrated that these correlations can significantly change EOP estimates. Schuh and Wilkin (1989) derived empirical correlation coefficients from 19 VLBI sessions, but did not take the next step of modifying the normal equations. Schuh and Tesmer (2000) derived empirical correlation coefficients, and, together with the a priori variance, σ^2_{meas}, constructed the covariance matrix. They demonstrated that this improved repeatability on 36 IRIS-S sessions from December 1994 through December 1998. Tesmer (2003a) (2003b), and Tesmer and Kutterer (2004) modified the covariance matrix by inflating the diagonal terms with additional contributions which were source, station and elevation dependent.

They found a reduction in the scatter of station position of a few percent. Gipson (2006) presented an earlier stage of this research at the 2006 IVS General Meeting.

In the next section, I present the least squares equations for VLBI. In Section 3 I discuss how station dependent delay modifies the covariance matrix. In Section 4, I study the effect of including two such error sources: "Clock-like" errors are observation independent; "Atmosphere-like" errors depend on the elevation. Using the CONT05 data set, I demonstrate that including these terms:

1. Decreases baseline scatter.
2. Gives more realistic formal errors.
3. Improves agreement between VLBI and GPS EOP estimates.

2 VLBI Normal Equations

The VLBI observable $\tau_{ij}(t)$ is the difference in arrival time of a signal at two stations i, j at time t. The delay is a function of various parameters A_a, some of which we are interested in, such as station position and Earth orientation, and other "nuisance" parameters such as clock drift and tropospheric delay.

Let $F_{a,ij}(t)$ be the derivative of the delay with respect to these parameters: $F_{a,ij}(t) \equiv \frac{\partial \tau_{ij}(t)}{\partial A_a}$. In the linear approximation the delay is:

$$\tau_{ij}(t) - \tau_{0,ij}(t) = \sum_a A_a F_{a,ij}(t) + \varepsilon_{ij,obs}(t)$$

Here $\tau_{0,ij}(t)$ is the a priori delay and $\varepsilon_{ij,obs}(t)$ is the noise associated with the observation. The noise term incorporates not only the correlator noise, but other sources of noise due to mis-calibration, mis-modeling, and failure of the linear approximation. This equation can be schematically re-written as:

$$\tau_{o-c} = FA + \varepsilon$$

Let Ω be the observation covariance matrix:

$$\Omega_{ijt,klt'} = < \varepsilon_{ij,obs}(t)\varepsilon_{kl,obs}(t') > \qquad (1)$$

Here the triples ijt and klt' label the observations. The least squares equations are given by:

$$\left(F^T\Omega^{-1}F\right)A = F^T\Omega^{-1}\tau_{o-c} \qquad (2)$$

where

$$F^T\Omega^{-1}F = \sum_{ijt}\sum_{klt'} F_{b,ij}(t)(t')\Omega^{-1}_{ijt,klt'}F_{a,kl}$$

and the sum is over all observations. There is a similar expression, *mutatis mutandis*, for $F^T\Omega^{-1}\tau_{o-c}$. The normal equations can formally be inverted to solve for A_a:

$$A = \left(F^T\Omega^{-1}F\right)^{-1}F^T\Omega^{-1}\tau_{o-c}$$

The usual assumption in VLBI data analysis is that the observations are independent, or, stated differently, the noise on different observations is uncorrelated. This is equivalent to saying that the covariance matrix is diagonal.

$$\Omega_{ijt,klt'} = \sigma^2_{ij}(t) \times \left(\delta^{ij}_{kl}\delta^{t}_{t'}\right)$$

In this case the normal equations simplify substantially, and we have, e.g.,

$$F^T\Omega^{-1}F = \sum_{ijt} F_{b,ij}(t)F_{a,ij}(t)\frac{1}{\sigma^2_{ij}(t)}$$

3 Effect of Station Dependent Delay Noise on Covariance Matrix

In this section I describe how to incorporate the effect of a particular kind of station dependent delay noise into the normal equations. Assume that the delay τ_i at station i is given by:

$$\tau_i = \tau_{i,geom} + \tau_{i,mod} + \varepsilon_{i,A} + \varepsilon_{i,B}\ldots.$$

$\tau_{i,geom}$ is the geometric delay in a vacuum, and $\tau_{i,mod}$ incorporates additional calibration and modeling terms. The $\varepsilon_{i,A}$ are station dependent delay error terms. The observational noise $\varepsilon_{ij,obs}(t)$ for baseline ij is:

$$\varepsilon_{ij,obs}(t) = \varepsilon_{ij,meas}(t) + \varepsilon_{ij,A}(t) + \varepsilon_{ij,B}(t) + \ldots$$

where $\varepsilon_{ij,meas}(t)$ is the measurement noise due to the correlation process, and the remaining terms are due to different kinds of station dependent delay error: $\varepsilon_{ij,A}(t) = \varepsilon_{i,A}(t) - \varepsilon_{j,A}(t)$.

The following assumptions simplify the evaluation of the covariance matrix:

1. The additional terms are uncorrelated with the measurement noise: $< \varepsilon_A \varepsilon_{meas} >= 0$
2. Different kinds of delay noise are uncorrelated: $< \varepsilon_A \varepsilon_B >= 0$ for $A \neq B$.

3. Delay errors at different times are uncorrelated: $< \varepsilon_{ij,A}(t)\varepsilon_{kl,A}(t') >= 0$ for $t \neq t'$.

4. Delay errors at different stations are uncorrelated.

In what follows, it is crucial to distinguish between *observations* and *scans*. An observation is the measurement of the delay on an *individual* baseline. A scan is a collection of *simultaneous* observations involving a commmon source.

By assumptions 1 and 2, the cross terms in the covariance matrix vanish:

$$\begin{aligned} \Omega &= < \varepsilon^2_{meas} > + < \varepsilon^2_A > + < \varepsilon^2_B > ... \\ &= \Omega_{meas} + \Omega_A + \Omega_B .... \end{aligned}$$

The first term Ω_{meas} is the (diagonal) covariance matrix associated with the measurement process. This is the the only kind of noise included in the standard analysis. The remaining terms are due to the additional noise sources.

Assumption 3 implies that the covariance matrix is block diagonal, with each block being the covariance matrix for a single scan.

By assumption 4, cross-terms involving different stations vanish. The diagonal elements for baseline ij are:

$$\begin{aligned} \Omega_{A,ij,ij} &= < (\varepsilon_{i,A} - \varepsilon_{j,A})^2 > \\ &= < \varepsilon^2_{i,A} + \varepsilon^2_{j,A} > \\ &= \sigma^2_{i,A} + \sigma^2_{j,A} \end{aligned}$$

i.e., just the sum of the noise terms for each station. This is what you would naively expect. The total diagonal contribution is

$$\begin{aligned} \Omega_{t,ij,ij} &= \sigma^2_{ij,meas} + \sigma^2_{i,A} + \sigma^2_{j,A} + \\ &\quad \sigma^2_{i,B} + \sigma^2_{j,B} + ... \end{aligned}$$

This increase in the diagonal terms of the covariance matrix increases the formal errors of the estimated parameters.

The station dependent noise terms also introduce off-diagonal terms in the covariance matrix which are non-zero *if, and only if* the baselines have a station in common. In this case we have:

$$\begin{aligned} \Omega_{A,ij,il} &= -\Omega_{A,ij,li} \\ &= < (\varepsilon_{i,A} - \varepsilon_{j,A})(\varepsilon_{i,A} - \varepsilon_{l,A}) > \\ &= < \varepsilon^2_{i,A} >= \sigma^2_{i,A} \end{aligned}$$

These off-diagonal terms also increase the formal erorrs of the estimates.

Note that both the diagonal and off-diagonal terms depend *only* on the variance of the noise. Hence station dependent delay noise has two effects: 1) The noise level of the observations is increased; and 2) Observations involving a common station at a given time are correlated. Both effects increase the formal erorrs, and both modify the VLBI estimates.

Since, by assumption 3, the covariance matrix is block diagonal, building up the normal equations given in Eq. (2) is straightforward and done on a scan by scan basis: 1) Compute the covariance matrix for a given scan; 2) Invert it; 3) Compute the contribution of this scan to the normal matrix.

4 Station Dependent Clock and Atmosphere Noise

I modified the VLBI analysis software *solve* to take into account two kinds of correlated station dependent noise. The first kind has a constant variance independent of the observation. This might be due, for example, to short term unmodeled clock variation, or random errors in the cable calibration. I call this kind of error "clock-like". Explicity, for station j the variance is:

$$\sigma^2_{j,clk} = a^2_{j,clk}$$

Another error source is due to mis-modeling the atmosphere. In this case, I assume that the variance takes the form:

$$\sigma^2_{j,atm} = a^2_{j,atm} \times [Map(el_j)]^2$$

where $Map(el_j)$ is the mapping function. Because of the strong elevation dependence of the mapping function, low-elevation points are downweigted. Both kinds of noise introduce correlations between observations.

In this study I make the simplifying assumption that the noise terms are the same for all stations, and do not vary with time. Realistically, of course, both noise sources may be time and station dependent.

I looked at the effect of including these noise sources on 15 CONT05 sessions. This is an example of a very good large network observing over a short period of time. Because the period is short, un-modeled seasonal effects should be small. Because the sessions were good, they should be more sensitive to improvements in the analysis.

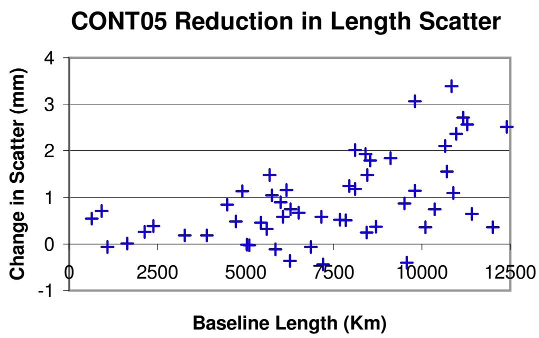

Figure 1. Difference in baseline repeatability for CONT05 data set between the standard solution and one incorporating 10 ps of atmosphere mapping function noise. Points above the x-axis are baselines where the scatter is reduced for the new solution.

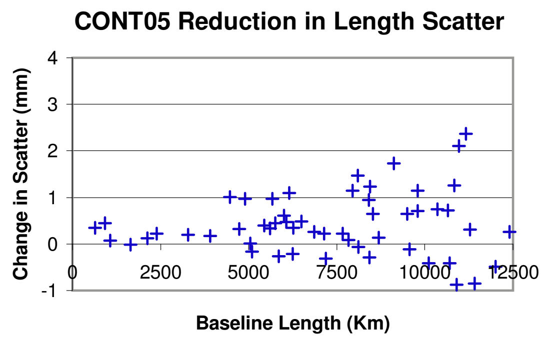

Figure 2. Difference in baseline repeatability for CONT05 data set between the standard solution and one incorporating 10 ps of atmosphere mapping function noise. Only the diagonal terms in the covariance matrix are used in the normal equations.

Figure 1 plots the difference in baseline scatter between the standard solution and a solution using the full covariance matrix where $a_{atm} = 10ps$. The new solution reduces the scatter for the vast majority (46 out of 54) baselines. The average improvement is 0.92 mm, or 12.4%. Generally speaking, the longer the baseline, the greater the improvement. On the few baselines where the scatter is worse, it is only slightly worse, i.e., well under a millimeter.

One possibility is that the improvement is due entirely or predominantly to the increase in the diagonal terms of the covariance matrix. To determine if this is true, I generated a parallel series of solutions where I included only the diagonal terms in the covariance matrix. Figure 2 is the baseline scatter plot for $a_{atm} = 10ps$. The scatter on most baselines (41/54) is improved, but the average improvement is 0.42 mm, less than half that obtained previously. This turns out to be a general feature: Including only the diagonal terms improves the solution, but not as much as including the full covariance matrix.

Table 1 summarizes the results of a series of solutions where I varied a_{atm} and a_{clk} independently. This table lists the results for using both the full covariance matrix and only the diagonal terms. For each pair of solutions the table displays: 1) Average χ^2; 2) RMS scatter; 3) Average improvement in millimeters; 4) Average improvement in percent; 5) Number of baselines where the scatter is reduced.

Introducing a clock-like noise source has little effect on the solutions. Introducing an atmosphere-like noise source generally improves the solutions with a broad peak starting around $a_{atm} = 10ps$. Note that in all instances, the results from the full covariance matrix are better than using just the diagonal terms: Including the correlations is beneficial. The χ^2 in this table is the average value over all baselines. Explictly, I use the formal errors for the baseline lengths given by *solve*, calculate the scatter and the χ^2 for each baseline, and then take the average over all baselines. For the standard solution, $\chi^2_{avg} = 2.16$, which indicates that the formal errors are too optimistic by $\sqrt{2.16} \simeq 1.47$. For the solution with $a_{atm} = 10ps$ $\chi^2_{avg} = 1.03$, indicating that the formal errors are, on average, correct. As the a_{atm} is increased beyond 10 ps, χ^2_{avg} decreases below 1, indicating that we have introduced too much extra noise.

Figure 3 plots the average improvement in baseline scatters a function of a_{atm}. It does this for both cases: full covariance matrix, or only diagonal terms. The greatest reduction in baseline scatter ocurrs around 10ps for both cases. For all values of a_{atm} the scatter is reduced, and for all values the reduction is greater if the full covariance matrix is used instead of the diagonal. Also plotted in this figure is the χ^2_{avg}. This starts at 2.16, and declines as the noise is increased. This is what you would expect. It is interesting to note that $\chi^2_{avg} \simeq 1$ at $a_{atm} = 10ps$: At the value where the baseline scatter has a mininum, the formal errors are, on average, correct.

The above examples show that including station dependent noise improves the consisteny of VLBI sessions. I also looked at the VLBI estimates of EOP, and compared them with GPS results. For this comparison I interpolated the GPS EOP estimates to the epoch of the VLBI estimates using a cubic spline. After removing

Table 1. Effect of Clock and Atmosphere Station Dependent Delay CONT05

		Full Covariance					Diagonal Only				
a_{clk}	a_{atm}	Avg	WRMS	Avg. Imp.		#BL	Avg	WRMS	Avg. Imp.		#BL
ps	ps	χ^2	mm	mm	%	Imp.	χ^2	mm	mm	%	Imp.
0	0	2.16	7.56	-	-	-	2.16	7.56	-	-	-
5	0	1.96	7.53	0.03	0.6%	30	2.07	7.57	-0.01	0.2%	26
10	0	1.65	7.52	0.04	1.0%	30	1.87	7.59	-0.03	0.1%	22
15	0	1.40	7.55	0.01	0.6%	28	1.67	7.65	-0.09	0.2%	21
20	0	1.21	7.62	-0.06	-0.4%	24	1.50	7.72	-0.16	-0.8%	20
25	0	1.07	7.72	-0.16	-1.7%	24	1.35	7.81	-0.25	-1.6%	20
0	1	2.09	7.48	0.08	1.3%	44	2.13	7.54	0.02	0.4%	36
0	2	1.93	7.32	0.24	3.8%	48	2.05	7.48	0.08	1.4%	42
0	3	1.76	7.16	0.40	6.1%	48	1.95	7.42	0.14	2.5%	43
0	4	1.61	7.02	0.54	7.9%	47	1.84	7.35	0.21	3.6%	43
0	5	1.47	6.91	0.65	9.3%	47	1.73	7.29	0.27	4.6%	43
0	6	1.36	6.83	0.73	10.4%	48	1.63	7.24	0.32	5.4%	43
0	7	1.25	6.76	0.80	11.2%	48	1.54	7.20	0.36	6.0%	41
0	8	1.17	6.70	0.86	11.8%	48	1.45	7.17	0.39	6.5%	42
0	9	1.09	6.66	0.90	12.1%	46	1.38	7.15	0.41	7.0%	41
0	10	1.03	6.64	0.92	12.4%	46	1.31	7.13	0.43	7.3%	41
0	11	0.97	6.62	0.94	12.5%	45	1.25	7.12	0.44	7.5%	40
0	12	0.92	6.60	0.96	12.6%	45	1.19	7.11	0.45	7.7%	40
0	13	0.87	6.60	0.96	12.5%	46	1.14	7.10	0.46	7.8%	40
0	14	0.83	6.60	0.96	12.4%	47	1.09	7.10	0.46	7.9%	39
0	15	0.79	6.61	0.95	12.2%	47	1.04	7.10	0.46	7.9%	39
0	20	0.65	6.68	0.88	11.0%	46	0.87	7.13	0.43	7.5%	36
0	25	0.55	6.79	0.77	9.4%	42	0.75	7.19	0.37	6.7%	36

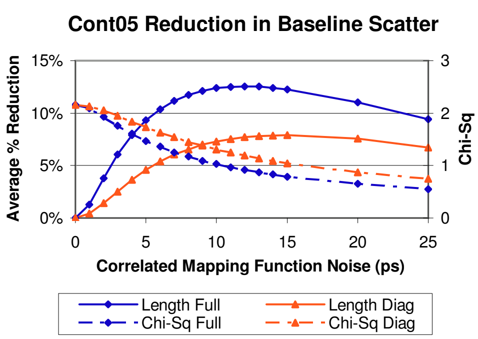

Figure 3. Average percentage improvement in baseline repeatability as a function of mapping function noise. Diamonds indicate the full covariance is used; triangles only the diagonal part of the covariance matrix. Solid lines give the average improvement in repeatability; dashed lines χ^2_{avg}.

an offset, I calculated the WRMS scatter of the difference for X-pole, Y-Pole, and LOD. Table 2 shows the results of this analysis. For each value of a_{atm} I display the WRMS scatter of the difference. For $a_{atm} > 0$, I also indicate the reduction in scatter compared to the $a_{atm} = 0$ solution. These results are presented graphically in Figure 4. As a_{atm} increases to 10 ps, the agreement between VLBI and GPS estimates of EOP improves. Beyond about 12ps there is relatively little change. Relatively speaking, the improvement in Y-Pole and LOD is much less than for X-Pole. The reasons for this are unclear.

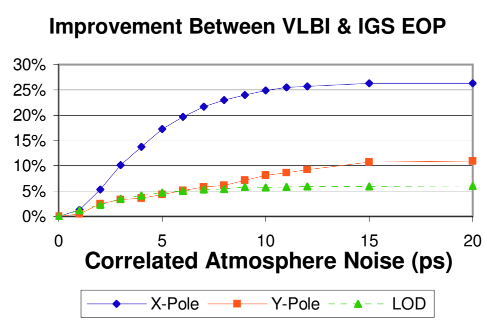

Figure 4. Including correlated atmosphere modeling error improves the agreement between VLBI and GPS estimates for EOP. Most of the improvement occurs in going from 0 ps to 10 ps of correlated noise. After 10 ps the improvement is modest.

Table 2. Comparison of VLBI & GPS EOP

	X-Pole		Y-Pole		LOD	
	RMS	%	RMS	%	RMS	%
a_{atm}	μas	Chg	μas	Chg	μs	Chg
0	60.4	-	39.2	-	17.0	-
1	59.6	1.3	39.0	0.5	16.8	1.2
2	57.2	5.3	38.2	2.6	16.6	2.4
3	54.3	10.1	37.9	3.3	16.4	3.5
4	52.1	13.7	37.8	3.6	16.3	4.1
5	50.0	17.2	37.5	4.3	16.2	4.7
6	48.5	19.7	37.2	5.1	16.1	5.1
7	47.3	21.7	36.9	5.9	16.1	5.4
8	46.5	23.0	36.8	6.1	16.1	5.4
9	45.9	24.0	36.4	7.1	16.0	5.8
10	45.4	24.8	36.0	8.2	16.0	5.8
11	45.0	25.5	35.8	8.7	16.0	5.8
12	44.9	25.7	35.6	9.2	16.0	5.8
15	44.5	26.3	35.0	10.7	16.0	5.8
20	44.5	26.3	34.9	11.0	16.0	5.8

5 Conclusion and Future Work

Using the CONT05 data set, I demonstrated that including station dependent delay noise reduces baseline scatter and results in more realistic formal errors. The effect of including clock-like errors is relatively small. In contrast, including the effect of atmospheric noise results in a dramatic decrease in baseline scatter. This improvement is not due simply to inflating the observational errors, but depends as well on the correlations introduced in the measurement. Including this error makes the VLBI estimates more consistent from day-to-day.

I also compared VLBI and GPS estimates of polar motion and LOD. I found that introducing correlated atmosphere noise improved the repeatability of all three components.

I have performed a similar analysis for other datasets including:

1. The complete set of R1s & R4s.
2. The RDV sessions.
3. High SNR experiments.
4. Sessions in which VLBI was measured by two simultaneous VLBI networks.

Lack of space prevents me from giving a full description of my results. In all instances including the effect of correlated noise reduces baseline scatter, gives more realistic formal errors, and improves the agreement between VLBI and GPS estimates of EOP. For the last data set it also improved the agreement between the VLBI estimates of EOP.

References

Gipson, J.M, (2006). Correlation due to Station Dependent Noise. 2006 IVS General Meeting Proceedings, Concepcion, Chile, D. Behrend and K. Baver (ed), pp 286-290, 2006.

Qian Zhi-han, (1985). The Correlation on VLBI Observables and Its Effects for the Determination of ERP. Shanghai Observatory.

Schuh, H. and Wilkin A. (1989) Determination of Correlation Coefficients between VLBI-Observables., Proc. Of the 7th Working Group Meeting on European VLBI, Madrid, Spain, CSIC, A. Rius Ed pp 79-91.

Schuh, H. & Tesmer, V: (2000) Considering A Priori Correlations in VLBI Data Analysis. 2000 IVS General Meeting Proceedings, p 237.

Tesmer, V. (2003a) Refinement of the Stochastic VLBI Model: First Results. 16th Working Meeting on European VLBI for Geodesy and Astrometry, pp 207-218.

Tesmer, V., Das stochastische Modell bei der VLBI-Anwertung, PhD thesis, University of Munich, Munich, Germany, 2003.

Tesmer, V., and Kutterer, H. (2004). An advanced Stochastic Model for VLBI Observations and its Applications to VLBI Data Analysis. 2004 IVS Generaml Meeting Proceedings, Ottawa, Canada, N. Vandenberg and K. Baver (ed), pp 296-300.